

IN THE SPECIFICATION

Please amend the title as follows:

--INFORMATION RETRIEVAL FROM A COLLECTION OF DATA--.

Please amend the paragraph beginning at page 5, line 6 as follows:

The normalized text is parsed 46, converting the normalized text into fragments adapted for further processing. Annotating words as putative [punitive] keys and values, according to a feature lexicon, produces fragments. The feature lexicon is a vocabulary, or book containing an alphabetical arrangement of the words in a language or of a considerable number of them, with the definition of each; a dictionary. For example, the feature lexicon may specify that the term "Compaq" is a potential value and that "CPU speed" is a potential key. Multiple annotations are possible.

Please amend the paragraph beginning at page 6, line 29 as follows:

If no valid interpretation exists, a determination 76 is made on whether the main database contains a valid interpretation. If there is a valid interpretation in the main database, the key value group is used 74. If no valid interpretation is found in the main database, the process 70 determines 78 whether previous index fields have a high confidence of uniquely containing the fragment. If so, the key value grouping is used 74. If not, other information sources are searched 80 and a valid key value group generated 82. If a high confidence and valid putative [punitive] key is determined through one of the information sources consulted, then the grouping of the key and value form an atomic element are used 74. To make it possible to override false interpretations, a configuration of grammar can also specify manual groupings of keys and values that take precedence over the meaning resolution process 70.

Please amend the paragraph beginning at page 18, line 20 as follows:

The information retrieval process 40 may also process a grammar and generate a grammar index, which can help find other phrased synonyms that other methods might not find. For example, "Xeon", an Intel Microprocessor whose full designation is the "Intel Pentium® Xeon Processor," may be represented in canonical form as "Intel Xeon Processor." If a user query is received for "Intel," "Xeon" would not be found without the grammar index of the

information access process 40. The information access process 40 will search the grammar index and produce a list of all grammar tokens containing "Intel," and add this list to the overall search so that the results would pick up "Xeon," among others.

Please amend the paragraph beginning at page 19, line 6 as follows:

Referring to FIG. 10, a query expansion process 250 includes normalizing 252 and parsing 254 the putative [punitive] text. The canonical non-terminal representations are inserted 256 into an IR index in place of the actual putative [punitive] text.

Please amend the paragraph beginning at page 19, line 10 amended as follows:

In an embodiment, the putative [punitive] text is used "as-is." However, when a search is requested by a user, the putative [punitive] search phrase is processed according to the grammar rules to obtain a canonical non-terminal representation. The grammar rules are then used in a generative manner to determine which other possible phrases could have generated the same canonical non-terminal representation. Those phrases are stored in the IR index.